

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number:

10/516,705

Source:

PCT

Date Processed by STIC:

12-20-04

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,705

TIME: 12:25:41

Input Set : A:\3056us0PSEQ.txt

Output Set: N:\CRF4\12202004\J516705.raw

```

3 <110> APPLICANT: HARA, Takahito
4     KUSAKA, Masami
5     MIYAZAKI, Junichi
7 <120> TITLE OF INVENTION: Mutated Androgen Receptor, Cancer Cells Expressing Same,
8     Production Method Thereof and Use Thereof
10 <130> FILE REFERENCE: 3056USOP
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/516,705
C--> 12 <141> CURRENT FILING DATE: 2004-12-02
12 <150> PRIOR APPLICATION NUMBER: PCT/JP03/06942
13 <151> PRIOR FILING DATE: 2003-06-02
15 <150> PRIOR APPLICATION NUMBER: JP 2002-162206
16 <151> PRIOR FILING DATE: 2002-06-03
18 <150> PRIOR APPLICATION NUMBER: JP 2002-255612
19 <151> PRIOR FILING DATE: 2002-08-30
21 <160> NUMBER OF SEQ ID NOS: 6
23 <170> SOFTWARE: PatentIn version 3.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 2775
27 <212> TYPE: DNA
28 <213> ORGANISM: Homo sapiens
30 <220> FEATURE:
31 <221> NAME/KEY: CDS
32 <222> LOCATION: (1)..(2775)
33 <223> OTHER INFORMATION:
W--> 35 <400> 1
36 atg gaa gtg cag tta ggg ctg gga agg gtc tac cct cgg ccg ccg tcc      48
37 Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser
38 1             5             10             15
39 aag acc tac cga gga gct ttc cag aat ctg ttc cag agc gtg cgc gaa      96
40 Lys Thr Tyr Arg Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu
41             20             25             30
42 gtg atc cag aac ccg ggc ccc agg cac cca gag gcc gcg agc gca gca      144
43 Val Ile Gln Asn Pro Gly Pro Arg His Pro Glu Ala Ala Ser Ala Ala
44             35             40             45
45 cct ccc ggc gcc agt ttg ctg ctg ctg cag cag cag cag cag cag cag      192
46 Pro Pro Gly Ala Ser Leu Leu Leu Leu Gln Gln Gln Gln Gln Gln Gln
47             50             55             60
48 cag cag cag cag cag cag cag cag cag cag cag cag cag cag cag      240
49 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
50 65             70             75             80
51 cag cag cag caa gag act agc ccc agg cag cag cag cag cag cag ggt      288
52 Gln Gln Gln Gln Glu Thr Ser Pro Arg Gln Gln Gln Gln Gln Gln Gly
53             85             90             95

```

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,705

TIME: 12:25:41

Input Set : A:\3056us0PSEQ.txt

Output Set: N:\CRF4\12202004\J516705.raw

54	gag	gat	ggt	tct	ccc	caa	gcc	cat	cgt	aga	ggc	ccc	aca	ggc	tac	ctg	336
55	Glu	Asp	Gly	Ser	Pro	Gln	Ala	His	Arg	Arg	Gly	Pro	Thr	Gly	Tyr	Leu	
56				100					105				110				
57	gtc	ctg	gat	gag	gaa	cag	caa	cct	tca	cag	ccg	cag	tcg	gcc	ctg	gag	384
58	Val	Leu	Asp	Glu	Glu	Gln	Gln	Pro	Ser	Gln	Pro	Gln	Ser	Ala	Leu	Glu	
59			115					120					125				
60	tgc	cac	ccc	gag	aga	ggt	tgc	gtc	cca	gag	cct	gga	gcc	gcc	gtg	gcc	432
61	Cys	His	Pro	Glu	Arg	Gly	Cys	Val	Pro	Glu	Pro	Gly	Ala	Ala	Val	Ala	
62		130					135					140					
63	gcc	agc	aag	ggg	ctg	ccg	cag	cag	ctg	cca	gca	cct	ccg	gac	gag	gat	480
64	Ala	Ser	Lys	Gly	Leu	Pro	Gln	Gln	Leu	Pro	Ala	Pro	Pro	Asp	Glu	Asp	
65	145				150				155				160				
66	gac	tca	gct	gcc	cca	tcc	acg	ttg	tcc	ctg	ctg	ggc	ccc	act	ttc	ccc	528
67	Asp	Ser	Ala	Ala	Pro	Ser	Thr	Leu	Ser	Leu	Leu	Gly	Pro	Thr	Phe	Pro	
68			165					170				175					
69	ggc	tta	agc	agc	tgc	tcc	gct	gac	ctt	aaa	gac	atc	ctg	agc	gag	gcc	576
70	Gly	Leu	Ser	Ser	Cys	Ser	Ala	Asp	Leu	Lys	Asp	Ile	Leu	Ser	Glu	Ala	
71			180					185				190					
72	agc	acc	atg	caa	ctc	ctt	cag	caa	cag	cag	cag	gaa	gca	gta	tcc	gaa	624
73	Ser	Thr	Met	Gln	Leu	Leu	Gln	Gln	Gln	Gln	Gln	Glu	Ala	Val	Ser	Glu	
74			195				200					205					
75	ggc	agc	agc	agc	ggg	aga	gcg	agg	gag	gcc	tcg	ggg	gct	ccc	act	tcc	672
76	Gly	Ser	Ser	Ser	Gly	Arg	Ala	Arg	Glu	Ala	Ser	Gly	Ala	Pro	Thr	Ser	
77		210				215			220								
78	tcc	aag	gac	aat	tac	tta	ggg	ggc	act	tcg	acc	att	tct	gac	aac	gcc	720
79	Ser	Lys	Asp	Asn	Tyr	Leu	Gly	Gly	Thr	Ser	Thr	Ile	Ser	Asp	Asn	Ala	
80	225				230				235				240				
81	aag	gag	ttg	tgt	aag	gca	gtg	tcg	gtg	tcc	atg	ggc	ctg	ggt	gtg	gag	768
82	Lys	Glu	Leu	Cys	Lys	Ala	Val	Ser	Val	Ser	Met	Gly	Leu	Gly	Val	Glu	
83			245					250				255					
84	gcg	ttg	gag	cat	ctg	agt	cca	ggg	gaa	cag	ctt	cgg	ggg	gat	tgc	atg	816
85	Ala	Leu	Glu	His	Leu	Ser	Pro	Gly	Glu	Gln	Leu	Arg	Gly	Asp	Cys	Met	
86			260				265				270						
87	tac	gcc	cca	ctt	ttg	gga	gtt	cca	ccc	gct	gtg	cgt	ccc	act	cct	tgt	864
88	Tyr	Ala	Pro	Leu	Leu	Gly	Val	Pro	Pro	Ala	Val	Arg	Pro	Thr	Pro	Cys	
89			275				280				285						
90	gcc	cca	ttg	gcc	gaa	tgc	aaa	ggt	tct	ctg	cta	gac	gac	agc	gca	ggc	912
91	Ala	Pro	Leu	Ala	Glu	Cys	Lys	Gly	Ser	Leu	Leu	Asp	Asp	Ser	Ala	Gly	
92		290				295			300								
93	aag	agc	act	gaa	gat	act	gct	gag	tat	tcc	cct	ttc	aag	gga	ggt	tac	960
94	Lys	Ser	Thr	Glu	Asp	Thr	Ala	Glu	Tyr	Ser	Pro	Phe	Lys	Gly	Gly	Tyr	
95	305				310				315				320				
96	acc	aaa	ggg	cta	gaa	ggc	gag	agc	cta	ggc	tgc	tct	ggc	agc	gct	gca	1008
97	Thr	Lys	Gly	Leu	Glu	Gly	Glu	Ser	Leu	Gly	Cys	Ser	Gly	Ser	Ala	Ala	
98			325				330				335						
99	gca	ggg	agc	tcc	ggg	aca	ctt	gaa	ctg	ccg	tct	acc	ctg	tct	ctc	tac	1056
100	Ala	Gly	Ser	Ser	Gly	Thr	Leu	Glu	Leu	Pro	Ser	Thr	Leu	Ser	Leu	Tyr	
101			340				345				350						
102	aag	tcc	gga	gca	ctg	gac	gag	gca	gct	gcg	tac	cag	agt	cgc	gac	tac	1104

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,705

TIME: 12:25:41

Input Set : A:\3056us0PSEQ.txt

Output Set: N:\CRF4\12202004\J516705.raw

103	Lys	Ser	Gly	Ala	Leu	Asp	Glu	Ala	Ala	Ala	Tyr	Gln	Ser	Arg	Asp	Tyr	
104			355					360					365				
105	tac	aac	ttt	cca	ctg	gct	ctg	gcc	gga	ccg	ccg	ccc	cct	ccg	ccg	cct	1152
106	Tyr	Asn	Phe	Pro	Leu	Ala	Leu	Ala	Gly	Pro	Pro	Pro	Pro	Pro	Pro	Pro	
107		370					375					380					
108	ccc	cat	ccc	cac	gct	cgc	atc	aag	ctg	gag	aac	ccg	ctg	gac	tac	ggc	1200
109	Pro	His	Pro	His	Ala	Arg	Ile	Lys	Leu	Glu	Asn	Pro	Leu	Asp	Tyr	Gly	
110	385					390				395					400		
111	agc	gcc	tgg	gcg	gct	gcg	gcg	gcg	cag	tgc	cgc	tat	ggg	gac	ctg	gcg	1248
112	Ser	Ala	Trp	Ala	Ala	Ala	Ala	Ala	Gln	Cys	Arg	Tyr	Gly	Asp	Leu	Ala	
113				405						410					415		
114	agc	ctg	cat	ggc	gcg	ggt	gca	gcg	gga	ccc	ggt	tct	ggg	tca	ccc	tca	1296
115	Ser	Leu	His	Gly	Ala	Gly	Ala	Ala	Gly	Pro	Gly	Ser	Gly	Ser	Pro	Ser	
116			420						425				430				
117	gcc	gcc	gct	tcc	tca	tcc	tgg	cac	act	ctc	ttc	aca	gcc	gaa	gaa	ggc	1344
118	Ala	Ala	Ala	Ser	Ser	Ser	Trp	His	Thr	Leu	Phe	Thr	Ala	Glu	Glu	Gly	
119		435					440					445					
120	cag	ttg	tat	gga	ccg	tgt	ggt	ggt	ggt	ggg	ggt	ggt	ggc	ggc	ggc	ggc	1392
121	Gln	Leu	Tyr	Gly	Pro	Cys	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	
122		450				455				460							
123	ggc	ggc	ggc	ggc	ggc	ggc	ggc	ggc	ggc	ggc	ggc	ggc	ggc	gag	gcg	gga	1440
124	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Glu	Ala	Gly	
125	465				470					475					480		
126	gct	gta	gcc	ccc	tac	ggc	tac	act	cgg	ccc	cct	cag	ggg	ctg	gcg	ggc	1488
127	Ala	Val	Ala	Pro	Tyr	Gly	Tyr	Thr	Arg	Pro	Pro	Gln	Gly	Leu	Ala	Gly	
128			485						490					495			
129	cag	gaa	agc	gac	ttc	acc	gca	cct	gat	gtg	tgg	tac	cct	ggc	ggc	atg	1536
130	Gln	Glu	Ser	Asp	Phe	Thr	Ala	Pro	Asp	Val	Trp	Tyr	Pro	Gly	Gly	Met	
131			500						505				510				
132	gtg	agc	aga	gtg	ccc	tat	ccc	agt	ccc	act	tgt	gtc	aaa	agc	gaa	atg	1584
133	Val	Ser	Arg	Val	Pro	Tyr	Pro	Ser	Pro	Thr	Cys	Val	Lys	Ser	Glu	Met	
134		515					520					525					
135	ggc	ccc	tgg	atg	gat	agc	tac	tcc	gga	cct	tac	ggg	gac	atg	cgt	ttg	1632
136	Gly	Pro	Trp	Met	Asp	Ser	Tyr	Ser	Gly	Pro	Tyr	Gly	Asp	Met	Arg	Leu	
137		530				535				540							
138	gag	act	gcc	agg	gac	cat	gtt	ttg	ccc	att	gac	tat	tac	ttt	cca	ccc	1680
139	Glu	Thr	Ala	Arg	Asp	His	Val	Leu	Pro	Ile	Asp	Tyr	Tyr	Phe	Pro	Pro	
140	545				550				555						560		
141	cag	aag	acc	tgc	ctg	atc	tgt	gga	gat	gaa	gct	tct	ggg	tgt	cac	tat	1728
142	Gln	Lys	Thr	Cys	Leu	Ile	Cys	Gly	Asp	Glu	Ala	Ser	Gly	Cys	His	Tyr	
143			565						570					575			
144	gga	gct	ctc	aca	tgt	gga	agc	tgc	aag	gtc	ttc	ttc	aaa	aga	gcc	gct	1776
145	Gly	Ala	Leu	Thr	Cys	Gly	Ser	Cys	Lys	Val	Phe	Phe	Lys	Arg	Ala	Ala	
146			580					585					590				
147	gaa	ggg	aaa	cag	aag	tac	ctg	tgc	gcc	agc	aga	aat	gat	tgc	act	att	1824
148	Glu	Gly	Lys	Gln	Lys	Tyr	Leu	Cys	Ala	Ser	Arg	Asn	Asp	Cys	Thr	Ile	
149			595				600					605					
150	gat	aaa	ttc	cga	agg	aaa	aat	tgt	cca	tct	tgt	cgt	ctt	cgg	aaa	tgt	1872
151	Asp	Lys	Phe	Arg	Arg	Lys	Asn	Cys	Pro	Ser	Cys	Arg	Leu	Arg	Lys	Cys	

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,705

TIME: 12:25:41

Input Set : A:\3056us0PSEQ.txt

Output Set: N:\CRF4\12202004\J516705.raw

152	610	615	620	
153	tat gaa gca ggg atg act ctg gga gcc cgg aag ctg aag aaa ctt ggt	1920		
154	Tyr Glu Ala Gly Met Thr Leu Gly Ala Arg Lys Leu Lys Lys Leu Gly			
155	625 630 635 640			
156	aat ctg aaa cta cag gag gaa gga gag gct tcc agc acc acc agc ccc	1968		
157	Asn Leu Lys Leu Gln Glu Glu Gly Glu Ala Ser Ser Thr Thr Ser Pro			
158	645 650 655			
159	act gag gag aca acc cag aag ctg aca gtg tca cac att gaa ggc tat	2016		
160	Thr Glu Glu Thr Thr Gln Lys Leu Thr Val Ser His Ile Glu Gly Tyr			
161	660 665 670			
162	gaa tgt cag ccc atc ttt ctg aat gtc ctg gaa gcc att gag cca ggt	2064		
163	Glu Cys Gln Pro Ile Phe Leu Asn Val Leu Glu Ala Ile Glu Pro Gly			
164	675 680 685			
165	gta gtg tgt gct gga cac gac aac aac cag ccc gac tcc ttt gca gcc	2112		
166	Val Val Cys Ala Gly His Asp Asn Asn Gln Pro Asp Ser Phe Ala Ala			
167	690 695 700			
168	ttg ctc tct agc ctc aat gaa ctg gga gag aga cag ctt gta cac gtg	2160		
169	Leu Leu Ser Ser Leu Asn Glu Leu Gly Glu Arg Gln Leu Val His Val			
170	705 710 715 720			
171	gtc aag tgg gcc aag gcc ttg cct ggc ttc cgc aac tta cac gtg gac	2208		
172	Val Lys Trp Ala Lys Ala Leu Pro Gly Phe Arg Asn Leu His Val Asp			
173	725 730 735			
174	gac cag atg gct gtc att cag tac tcc tgg atg ggg ctc atg gtg ttt	2256		
175	Asp Gln Met Ala Val Ile Gln Tyr Ser Trp Met Gly Leu Met Val Phe			
176	740 745 750			
177	gcc atg ggc tgg cga tcc ttc acc aat gtc aac tcc agg atg ctc tac	2304		
178	Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser Arg Met Leu Tyr			
179	755 760 765			
180	ttc gcc cct gat ctg gtt ttc aat gag tac cgc atg cac aag tcc cgg	2352		
181	Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met His Lys Ser Arg			
182	770 775 780			
183	atg tac agc cag tgt gtc cga atg agg cac ctc tct caa gag ttt gga	2400		
184	Met Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser Gln Glu Phe Gly			
185	785 790 795 800			
186	tgg ctc caa atc acc ccc cag gaa ttc ctg tgc atg aaa gca ctg cta	2448		
187	Trp Leu Gln Ile Thr Pro Gln Glu Phe Leu Cys Met Lys Ala Leu Leu			
188	805 810 815			
189	ctc ttc agc att att cca gtg gat ggg ctg aaa aat caa aaa ttc ttt	2496		
190	Leu Phe Ser Ile Ile Pro Val Asp Gly Leu Lys Asn Gln Lys Phe Phe			
191	820 825 830			
192	gat gaa ctt cga atg aac tac atc aag gaa ctc gat cgt atc att gca	2544		
193	Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp Arg Ile Ile Ala			
194	835 840 845			
195	tgc aaa aga aaa aat ccc aca tcc tgc tca aga cgc ttc tac cag ctc	2592		
196	Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Phe Tyr Gln Leu			
197	850 855 860			
198	acc aag ctc ctg gac tcc gtg cag cct att gcg aga gag ctg cat cag	2640		
199	Thr Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln			
200	865 870 875 880			

## RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,705

TIME: 12:25:41

Input Set : A:\3056us0PSEQ.txt

Output Set: N:\CRF4\12202004\J516705.raw

```

201 ttc act ttt gac ctg cta atc aag tca cac atg gtg agc gtg gac ttt      2688
202 Phe Thr Phe Asp Leu Leu Ile Lys Ser His Met Val Ser Val Asp Phe
203      885      890      895
204 ccg gaa atg atg gca gag atc atc tct gtg caa gtg ccc aag atc ctt      2736
205 Pro Glu Met Met Ala Glu Ile Ile Ser Val Gln Val Pro Lys Ile Leu
206      900      905      910
207 tct ggg aaa gtc aag ccc atc tat ttc cac acc cag tga      2775
208 Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln
209      915      920
212 <210> SEQ ID NO: 2
213 <211> LENGTH: 924
214 <212> TYPE: PRT
215 <213> ORGANISM: Homo sapiens
217 <400> SEQUENCE: 2
218 Met Glu Val Gln Leu Gly Leu Gly Arg Val Tyr Pro Arg Pro Pro Ser
219 1      5      10      15
220 Lys Thr Tyr Arg Gly Ala Phe Gln Asn Leu Phe Gln Ser Val Arg Glu
221      20      25      30
222 Val Ile Gln Asn Pro Gly Pro Arg His Pro Glu Ala Ala Ser Ala Ala
223      35      40      45
224 Pro Pro Gly Ala Ser Leu Leu Leu Leu Gln Gln Gln Gln Gln Gln Gln
225      50      55      60
226 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
227 65      70      75      80
228 Gln Gln Gln Gln Glu Thr Ser Pro Arg Gln Gln Gln Gln Gln Gln Gly
229      85      90      95
230 Glu Asp Gly Ser Pro Gln Ala His Arg Arg Gly Pro Thr Gly Tyr Leu
231      100      105      110
232 Val Leu Asp Glu Glu Gln Gln Pro Ser Gln Pro Gln Ser Ala Leu Glu
233      115      120      125
234 Cys His Pro Glu Arg Gly Cys Val Pro Glu Pro Gly Ala Ala Val Ala
235      130      135      140
236 Ala Ser Lys Gly Leu Pro Gln Gln Leu Pro Ala Pro Pro Asp Glu Asp
237 145      150      155      160
238 Asp Ser Ala Ala Pro Ser Thr Leu Ser Leu Leu Gly Pro Thr Phe Pro
239      165      170      175
240 Gly Leu Ser Ser Cys Ser Ala Asp Leu Lys Asp Ile Leu Ser Glu Ala
241      180      185      190
242 Ser Thr Met Gln Leu Leu Gln Gln Gln Gln Gln Glu Ala Val Ser Glu
243      195      200      205
244 Gly Ser Ser Ser Gly Arg Ala Arg Glu Ala Ser Gly Ala Pro Thr Ser
245      210      215      220
246 Ser Lys Asp Asn Tyr Leu Gly Gly Thr Ser Thr Ile Ser Asp Asn Ala
247 225      230      235      240
248 Lys Glu Leu Cys Lys Ala Val Ser Val Ser Met Gly Leu Gly Val Glu
249      245      250      255
250 Ala Leu Glu His Leu Ser Pro Gly Glu Gln Leu Arg Gly Asp Cys Met
251      260      265      270
252 Tyr Ala Pro Leu Leu Gly Val Pro Pro Ala Val Arg Pro Thr Pro Cys

```

## VERIFICATION SUMMARY

DATE: 12/20/2004

PATENT APPLICATION: US/10/516,705

TIME: 12:25:42

Input Set : A:\3056us0PSEQ.txt

Output Set: N:\CRF4\12202004\J516705.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:35 M:258 W: Mandatory Feature missing, &lt;223&gt; Blank for SEQ#:1,Line#:33